Study Report 97-05

Adverse Impact Implications of Selection Instrument Group Score Differences

Jay M. Silva
U.S. Army Research Institute

19980311 048

United States Army Research Institute
for the Behavioral and Social Sciences

March 1997

U.S. ARMY RESEARCH INSTITUTE
FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency Under the Jurisdiction of the Deputy Chief of Staff for Personnel

EDGAR M. JOHNSON
Director

Technical review by

Peter Greenston
Leonard White

REPORT DOCUMENTATION PAGE

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED (from. . . to)
1997, March	Final	March 1995-June 1996

4. TITLE AND SUBTITLE Adverse Impact Implications of Selection Instrument Group Score Differences	5a. CONTRACT OR GRANT NUMBER 5b. PROGRAM ELEMENT NUMBER 0605803A
6. AUTHOR(S) Jay M. Silva	5c. PROJECT NUMBER D730 5d. TASK NUMBER 1231 5e. WORK UNIT NUMBER H01
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) U.S. Army Research Institute for the Behavioral and Social Sciences ATTN: PERI-RS 5001 Eisenhower Avenue Alexandria, VA 22333-5600	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Research Institute for the Behavioral and Social Sciences 5001 Eisenhower Avenue Alexandria, VA 22333-5600	10. MONITOR ACRONYM ARI 11. MONITOR REPORT NUMBER Study Report 97-05

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT *(Maximum 200 words)*:

Human resources decision-makers are concerned when mean inter-group score differences on selection measures are observed. Moreover, they are not concerned with the magnitude of the differences *per se*, but rather with whether those score differences will manifest themselves as adverse impact. An analytical approach was used to estimate for various combinations of selection ratio and minority applicant group representation, the maximum group score difference that would not violate the "four-fifths" rule. In addition, applicant pools of specific sizes with no mean inter-group score difference on the selection measure were considered to compute the conservative likelihood of encountering an adverse impact situation in a specific applicant sample. The results clearly suggest that the identification of adverse impact and its statistical substantiation will often occur in small applicant pools (i.e., 100), even when there is a small inter-group difference on the selection measure. For larger samples (i.e., 500), the results suggest that adverse impact will often be indicated when small mean inter-group selection measure differences are present. It is not clear to what degree the adverse impact found would be statistically substantiated. Research focusing on adverse impact and its statistical substantiation is needed for specific inter-group difference/applicant pool size combinations to create a clearer equivalence between inter-group differences and adverse impact.

15. SUBJECT TERMS

Adverse impact | Group differences | "Four-fifths" rule | Selection

SECURITY CLASSIFICATION OF			19. LIMITATION OF ABSTRACT	20. NUMBER OF PAGES	21. RESPONSIBLE PERSON (Name and Telephone Number)
16. REPORT	17. ABSTRACT	18. THIS PAGE			
Unclassified	Unclassified	Unclassified	Unlimited	21	

Study Report 97-05

Adverse Impact Implications of Selection Instrument Group Score Differences

Jay M. Silva
U.S. Army Research Institute

Selection and Assignment Research Unit
Michael G. Rumsey, Chief

U.S. Army Research Institute for the Behavioral and Social Sciences
5001 Eisenhower Avenue, Alexandria, Virginia 22333-5600

Office, Deputy Chief of Staff for Personnel
Department of the Army

March 1997

Army Project Number
20465803D730

Personnel and Training
Analysis Activities

FOREWORD

The presence of group differences creates many problems for those involved in selection testing. Among the problems it creates is that the impact on simple minority representation, and its more complex counterpart in the form of adverse impact, is not clear in traditional metrics.

This research takes a large step forward in translating an observed group difference to its impact on minority representation. It will enable researchers to communicate more clearly about group differences and their expected impact, and to place a greater focus on the direct needs of decision-makers.

ZITA M. SIMUTIS	EDGAR M. JOHNSON
Technical Director	Director

ACKNOWLEDGMENTS

I gratefully acknowledge the editorial assistance of Fred Mael on an earlier draft of this paper.

ADVERSE IMPACT IMPLICATIONS OF SELECTION INSTRUMENT GROUP SCORE DIFFERENCES

EXECUTIVE SUMMARY

Research Requirement:

Convey information on group differences in selection measures in a way that more directly clarifies when a given group difference would be expected to lead to a finding of adverse impact as defined by the "four-fifths" rule outlined in the Uniform Guidelines on Employment Selection Procedures (1978). In addition, because applicant pools of limited size, rather than infinite size samples used in computing expected means, are the practical focus, estimate for the conservative case of no inter-group difference, the likelihood of identifying adverse impact, and statistically substantiating it in samples of 100, 500, and 5,000 applicants.

Procedure:

Analytical formulas were developed to allow direct evaluation of when a group difference would be expected to lead to the identification of adverse impact. An analytical approach was also developed to compute the probability of identifying adverse impact, the probability of statistically substantiating it, and the expected value of the ratio of the selection rations within two groups and its variability for applicant pools of 100, 500, and 5,000 applicants. All analyses varied the combination of selection ratio and minority applicant pool representation.

Findings:

For typical situations involving a selection ratio of .50 or less, the results clearly showed that small-group differences of 0.10 to 0.27 of a standard deviation are all that can be tolerated before one would expect to be faced with an adverse impact problem. When applied to specific applicant samples it is clear that even a mean inter-group difference of 0.10 of a standard deviation would be problematic when the applicant pool is only 100. This can be extrapolated from the fact that adverse impact is problematic with applicant pools of 100, even when there is no inter-group difference on the selection measure. For applicant pools of 500 with no mean inter-group difference, an adverse impact finding is still highly possible, although it is not likely to be substantiated statistically. It is not known what the degree of the problem would be when the mean inter-group difference on the selection measure rose to 0.10 or higher for applicant pools of 500, but the probability of finding adverse impact would be higher as would the probability of statistically substantiating it.

Utilization of Findings:

The approach used to quantify group differences clarifies the degree to which one is likely to violate legal standards of minority representation when using a selection instrument exhibiting mean group differences. It is easily understood by those in the selection testing community and is directly applicable to selection decisions.

ADVERSE IMPACT IMPLICATIONS OF SELECTION INSTRUMENT GROUP SCORE DIFFERENCES

CONTENTS

LIST OF TABLES

ADVERSE IMPACT IMPLICATIONS OF SELECTION INSTRUMENT GROUP SCORE DIFFERENCES

INTRODUCTION

Representing group performance differences in the standard deviation metric is appropriate and commonly done in studies which examine inter-group performance differences (Coleman et al, 1966; Grant & Bray, 1970; Hunter, 1983; Hunter, Schmidt, & Rauschenberger, 1977; Hyde, Fennema, & Lamon, 1990; Hyde & Linn, 1988). The standard deviation metric informs on the magnitude of the group performance difference while maintaining an interval scale. But while the standard deviation metric is an appropriate metric for researchers, it is not a meaningful one for human resources decision-makers with only a cursory acquaintance with statistics. When communicating group differences in standard deviation units to these individuals, their lack of familiarity with the terminology and the normal distribution prevents them from understanding the magnitude and importance (or irrelevance) of the difference.

For these decision-makers it would be more practical to convert the standard deviation group difference to a metric which they can use to directly assess the viability of a selection instrument in their organization. One way that many organizations assess the viability of a selection instrument is to estimate whether when using it they would likely violate the "four-fifths" rule. The "four-fifths" rule as defined in the Uniform Guidelines on Employment Selection Procedures (1978) is violated when a selection rate for any racial, ethnic, or sex subgroup is less than "four-fifths" (4/5 or eighty percent) of the rate for the group with the highest rate. Additionally, the Guidelines indicate that a violation of the "four-fifths" rule "will generally be regarded as evidence of adverse impact . . ." Since the "four-fifths" rule is used to establish a prima facie case of adverse impact it is a point of concern for employers who are concerned with the litigation costs to defend a valid, but adverse-impact producing, selection instrument.

Human resources decision-makers are aware that when there is an average group score difference indicating lower average scores for minority groups such as Blacks, Hispanics, and females, there is a higher likelihood that the rates of minority hiring will be lower than those expected under the "four-fifths" rule. However, they are not able to determine whether a specific average group score difference would tend to violate the "four-fifths" rule in their organizational context (i.e., the selection ratio and minority applicant percentage typically found in their environment). The same is true for I/O psychologists. There is no simple way to use the normal distribution table to determine the maximum inter-group difference which is not expected to violate the "four-fifths" rule. There are a variety of reasons for this including the need to consider two normal distributions, and the impact of the selection ratio and minority applicant base rate.

In addition, it is hypothesized that even small group differences, when applied to a specific small group of applicants, would make it highly likely that the "four-fifths" rule would be violated. Thus, to illustrate just how likely it is to violate the "four-fifths" rule in specific organizational contexts with specific applicant numbers (i.e., 100, 500, and 5,000), the probability of violating the four-fifths rule was examined for the case where there was no mean group difference on the selection measure. When there are group differences, the probability of a violation should increase substantially. In addition, this paper also examines the probability of determining that the

"four-fifths" rule violation is statistically significant if a statistical test is conducted after identifying the violation when there is no mean group difference on the selection measure.

This paper will provide I/O psychologists with the means to inform decision-makers on whether the "four-fifths" rule is likely to be violated as the result of making hiring decisions based on selection test scores manifesting a specific average group score difference. The results, based on specific sized samples with no mean group difference on the selection measure, will clarify the minimum likelihood of a "four-fifths" rule violation, and subsequent likelihood that the violation is statistically significant upon further examination.

METHOD

The applied analytical approaches assumed the following: a) two groups of individuals (i.e., a lower-scoring group and a higher-scoring group), b) for each group the distribution of test scores was normal, c) the standard deviation of the test scores was equal across the two groups and thus could be expressed in standard score form (i.e., $\sigma_{lower} = \sigma_{higher} = 1$), d) the mean test score was scaled to equal zero for the higher-scoring group, while the mean test score for the lower-scoring group was presumed to be lower by delta (Δ), e) selection was accomplished using a single-list of test scores in a top-down fashion, and f) all selected applicants accepted the employment offer.

Computing the Maximum Inter-group Mean Score Difference Which Does Not Violate the "Four-fifths" Rule

Initially delta was set to a value of 0.01 concurrently with setting and fixing the selection ratio and lower-scoring group applicant base rate (i.e., proportion of applicants who are lower-performing group members) to specific values, a and b, respectively. The expected within-group proportion of higher-performing group applicants selected was computed as:

$$E\,[\,p_{higher}\,] = \int_{c}^{+\infty} f(x)\;dx\,. \tag{1}$$

Where c is the test standard score cutoff for selection. Since no analytical solution exists to compute c, Newton's iterative method was used as demonstrated in Hunter et al. (1977). The cutoff c is determined by the selection ratio, the lower-scoring group applicant base rate, and the mean expected score difference between the lower- and highest-scoring groups (delta). The involvement of the selection ratio and the lower-scoring group applicant base rate in determining c is the reason why these values were fixed at values a and b, respectively.

Similarly, the expected within-group proportion of lower-performing group applicants selected was computed as:

$$E\,[\,p_{lower}\,] = \int_{c+\Delta}^{+\infty} f(x)\;dx\,. \tag{2}$$

Next, the value from equation (2) was divided by the value from equation (1). If the value of this ratio was greater than .80 then delta was increased by an increment of 0.01 and the procedure was repeated. This procedure was repeated until the ratio of the value from equation (2) over the value from equation (1) fell below .80. At that point delta was decreased by 0.01 and this delta value was the maximum inter-group difference for the selection ratio value of a and the lower-scoring group applicant base rate value of b.

Independent Variables. Three independent variables were specified: a) the selection ratio, b) the lower-scoring group applicant base rate, and c) the mean expected score difference between the lower- and highest-scoring groups. The selection ratio was varied from .05 to .95 in intervals of .05 and the lower-scoring group applicant base rate was varied from .05 to .50 in intervals of .05. Finally, the mean expected score difference between the lower- and highest-scoring groups was initially set at 0.01 standard deviations and allowed to attain a value as high as 9.00 in increments of 0.01. It will be noted here that aspects of selection related to the criterion such as inter-group validity differences and inter-group criterion performance differences have no impact whatsoever on the determination of the maximum delta which will not violate the "four-fifths" rule since the "four-fifths" rule is wholly concerned with predictor effects.

Dependent Variable. For each combination of selection ratio and lower-scoring group applicant base rate the maximum inter-group mean score difference (delta) which was not expected to violate the "four-fifths" rule was determined.

Computing the Probability of Violating the "Four-fifths" Rule with a Specific Number of Applicants when the Groups Perform Equally Well on the Selection Measure

Independent Variables. Three independent variables were specified: a) the selection ratio, b) the lower-scoring group applicant base rate, and c) the number of applicants. The selection ratio was varied from .10 to .90 in intervals of .10, the lower-scoring group applicant base rate was varied from .10 to .50 in intervals of .10, and the number of applicants were 100, 500, and 5,000. The mean expected score difference between the lower- and highest-scoring groups (Δ) was zero.

Dependent Variables. Four dependent variables were computed: a) the probability of a "four-fifths" rule violation in the presence of no group differences, b) the probability that a "four-fifths" rule violation would be found to be statistically significant, c) the expected ratio of minority to majority ratios, and d) the standard deviation of the expected ratio.

The approach for computing the probability that the "four-fifths" rule will be violated when a specific number of applicants is encountered is quite different from the approach described above. First, given the number of applicants to be selected (n; number of total applicants multiplied by the selection ratio) and the proportion of applicants which are minority (p_{min}; number of minority applicants divided by the number of total applicants), all the possible minority/majority hiring combinations are determined (i.e., in terms of m minority applicants hired; where m varies between 0 and number of minority applicants or number of hires, whichever is lowest). Second, the probability of each hiring combination is computed as follows:

$$p_{combination} = \sum_{j=0}^{m} \binom{n}{j} p_{\min}^{j}(1 - p_{\min})^{n-j} - \sum_{j=0}^{m-1} \binom{n}{j} p_{\min}^{j}(1 - p_{\min})^{n-j} \qquad when\ m \geq 1 \qquad \textbf{(3)}$$

$$p_{combination} = \binom{n}{0} p_{\min}^{j}(1 - p_{\min})^{n-j} \qquad when\ m=0 \qquad \textbf{(4)}$$

Formulas 3 and 4 represent sampling m individuals from a pool of n individuals without replacement. $P_{combination}$ is the probability of hiring m minority applicants from a fixed pool of applicants without replacement. Third, for each combination determine if it is a violation of the "four-fifths" rule, and if it is then sum the probability of the combination occurring to a running probability total indicating the likelihood of violating the "four-fifths" rule.

In addition to determining the probability of a "four-fifths" rule violation it is also possible to easily compute some additional useful statistics. First, for each combination of minority and majority hires which violates the "four-fifths" rule determine if the violation is statistically significant, and if so sum the probability of the combination occurring to a running probability total indicating the likelihood that a statistical test would find the combination of minority and majority hires to be a statistically significant violation of the "four-fifths" rule. Divide this probability by the probability of violating the "four-fifths" rule to determine the probability of finding statistical significance when a violation is found.

Second, for each combination determine the ratio of the hiring ratios within each group and multiply it by the probability of the combination of minority and majority hires, and sum it to a running total indicating the mean ratio of the hiring proportions (i.e., for each group).

Finally, it is possible to compute the standard deviation of the mean ratio by maintaining a running total of the squared deviation of each combination's ratio from the overall mean multiplied by the probability of each combination, and then square rooting the running total.

RESULTS

Table 1 presents the maximum inter-group mean score difference (delta) which was not expected to violate the "four-fifths" rule, as a function of the selection ratio and lower-scoring group applicant base rate. At lower selection ratios the maximum delta was lower, and the lower-scoring group applicant base rate made little if any impact. However, at higher selection ratios (i.e., above .70), higher applicant base rates for the lower-scoring group increased the value of the inter-group mean score difference possible before the "four-fifths" rule was violated. Finally, as

the selection ratio increased, the maximum delta progressively increased in an accelerated fashion which was most noticeable at selection ratios above .70.

For a typical selection ratio of .50 and lower-scoring group applicant base rate of .20, the standard deviation inter-group score difference would have to exceed 0.26 standard deviations before the "four-fifths" rule would be violated. This means that, on average, an inter-group score difference of 0.26 standard deviations on a selection instrument will not violate the "four-fifths" rule. The "on average" component is to indicate that the "four-fifths" rule may be violated for any one sample of applicants as a result of random error in the sampling. However, in the long run (i.e., an infinite number of applicants) the "four-fifths" rule would not be violated.

One should be most concerned about violating the "four-fifths" rule when the selection ratio is low (i.e., below .30). At more moderate selection ratios (i.e., at least .50) it is expected that the "four-fifths" rule would not be violated at inter-group score differences below 0.26 standard deviations. Finally, when the selection ratios are even higher (i.e., equal or greater than .80) it is expected that "four-fifths" rule would not be violated at inter-group score differences below 0.50 standard deviations.

Tables 2, 3, and 4 address the issue of the probability of a "four-fifths" rule violation for specific numbers of applicants (i.e., 100, 500, an 5,000). Since for the case examined the two groups exhibited no mean difference on the selection measure, the numbers in these tables should be viewed as the minimum probability that a violation would be observed. After examining these three tables, it was quite clear and not surprising that a violation is most likely to be identified when the number of applicants is smaller. However, it was surprising that with 100 applicants the likelihood of violating the "four-fifths" rule was as high as shown (i.e.,as high as 43% of the time we would expect to violate the "four-fifths" rule). The good news was that upon exploring the statistical significance of an individual violation we would not be likely to find that the violation was statistically meaningful for selection ratios below .70. For selection ratios above .70 the likelihood of associating statistical significance with the violation increases to a maximum of .35 when the selection ratio is .90 and the minority applicant base rate is .20.

In addition, the expected ratio of the hiring rates for an applicant pool of 100 varied substantially, and in some situations (i.e., high selection ratio) was below .80. The degree of variability (i.e., standard deviation) around the expected ratio also varied substantially, ranging from 0.21 to 1.23, demonstrating just how violation volatile the situation is with 100 applicants.

When the situation shifts to a 500 member applicant pool, the situation is still surprisingly volatile. The probability that the "four-fifths" rule would be violated was still as high as 43% of the time. The good news was that with 500 applicants the probability of a finding statistical significance for an individual violation would be in line with acceptable error rates. The probability would be no higher than .05 for all conditions except a selection ratio of .90 with accompanying low minority representation in the applicant pool. In addition, the expected ratio of the hiring rates for an applicant pool of 500 varied much less, hovering around 1.00 with lower

Table 1

Maximum Inter-Group Mean Score Difference Which Does not Violate the Four-Fifths Rule As a Function of the Selection Ratio and Lower-Performing Group Applicant Base Rate

Selection Ratio	Lower-Performing Group Applicant Base Rate									
	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0.05	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10
0.10	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12
0.15	0.13	0.13	0.13	0.14	0.14	0.14	0.14	0.14	0.14	0.14
0.20	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15
0.25	0.16	0.16	0.16	0.17	0.17	0.17	0.17	0.17	0.17	0.17
0.30	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0.19	0.19
0.35	0.19	0.19	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.21
0.40	0.21	0.21	0.21	0.22	0.22	0.22	0.22	0.22	0.22	0.23
0.45	0.23	0.23	0.23	0.24	0.24	0.24	0.24	0.24	0.25	0.25
0.50	0.25	0.25	0.26	0.26	0.26	0.26	0.27	0.27	0.27	0.27
0.55	0.27	0.28	0.28	0.28	0.29	0.29	0.29	0.30	0.30	0.31
0.60	0.30	0.31	0.31	0.31	0.32	0.32	0.33	0.33	0.34	0.34
0.65	0.33	0.34	0.34	0.35	0.36	0.36	0.37	0.37	0.38	0.39
0.70	0.37	0.38	0.39	0.39	0.40	0.41	0.42	0.43	0.44	0.45
0.75	0.42	0.43	0.44	0.45	0.46	0.48	0.49	0.50	0.52	0.53
0.80	0.49	0.50	0.52	0.53	0.55	0.57	0.59	0.61	0.63	0.66
0.85	0.58	0.60	0.62	0.65	0.68	0.71	0.74	0.79	0.84	0.90
0.90	0.73	0.76	0.80	0.85	0.92	0.99	1.09	1.23	1.48	
0.95	1.01	1.11	1.25	1.49						

Notes. Maximum inter-group mean score difference in body of table are in standard deviation units. A blank cell indicates that regardless of the size of the inter-group mean score difference, the "four-fifths" rule can never be violated.

Table 2

Proportion of Time Organization Will Violate Four-Fifths Rule, the Probability of Identifying the Violation as Statistically Significant, and the Mean Minority/Majority Applicant Hiring Ratio and Its Standard Deviation with 100 Applicants

Selection Ratio	Minority Group Applicant Base Rate				
	0.10	0.20	0.30	0.40	0.50
0.10	0.35[a] 0.00[b] 1.13[c] (1.23)[d]	0.38 0.00 1.15 (0.98)	0.38 0.00 1.18 (0.95)	0.38 0.00 1.23 (1.03)	0.38 0.00 1.29 (1.14)
0.20	0.39 0.00 1.06 (0.80)	0.41 0.00 1.07 (0.62)	0.42 0.00 1.08 (0.56)	0.25 0.00 1.10 (0.55)	0.25 0.00 1.12 (0.59)
0.30	0.41 0.00 1.04 (0.64)	0.26 0.00 1.04 (0.49)	0.28 0.00 1.05 (0.43)	0.29 0.00 1.06 (0.42)	0.29 0.00 1.07 (0.43)
0.40	0.42 0.00 1.02 (0.54)	0.29 0.00 1.03 (0.41)	0.31 0.00 1.04 (0.37)	0.21 0.00 1.04 (0.35)	0.21 0.01 1.05 (0.35)
0.50	0.43 0.01 1.00 (0.46)	0.31 0.00 1.02 (0.36)	0.22 0.01 1.03 (0.32)	0.24 0.01 1.04 (0.31)	0.24 0.01 1.04 (0.31)
0.60	0.27 0.05 0.94 (0.38)	0.21 0.02 1.01 (0.32)	0.24 0.02 1.02 (0.29)	0.18 0.03 1.03 (0.28)	0.18 0.04 1.04 (0.28)
0.70	0.29 0.08 0.85 (0.32)	0.23 0.09 0.96 (0.27)	0.18 0.06 1.01 (0.26)	0.20 0.05 1.02 (0.25)	0.20 0.06 1.03 (0.25)
0.80	0.30 0.12 0.73 (0.28)	0.25 0.12 0.85 (0.23)	0.20 0.15 0.92 (0.21)	0.15 0.15 0.98 (0.21)	0.15 0.13 1.00 (0.22)
0.90	0.31 0.33 0.59 (0.27)	0.17 0.35 0.67 (0.25)	0.11 0.23 0.71 (0.24)	0.09 0.00 0.74 (0.24)	0.00 0.00 0.76 (0.24)

Notes. [a] The probability of violating the "four-fifths" rule.
[b] The probability that a "four-fifths" rule violation will be statistically significant (i.e., the ratio is statistically less than .80).
[c] The expected ratio (i.e., mean).
[d] The standard deviation of the expected ratio.

Table 3

Proportion of Time Organization Will Violate Four-Fifths Rule, the Probability of Identifying the Violation as Statistically Significant, and the Mean Minority/Majority Applicant Hiring Ratio and Its Standard Deviation with 500 Applicants

Selection Ratio	Minority Group Applicant Base Rate				
	0.10	0.20	0.30	0.40	0.50
0.10	0.43[a] 0.00[b] 1.02[c] (0.48)[d]	0.31 0.00 1.03 (0.37)	0.22 0.00 1.03 (0.32)	0.24 0.00 1.04 (0.31)	0.24 0.00 1.04 (0.31)
0.20	0.32 0.00 1.01 (0.34)	0.19 0.00 1.01 (0.25)	0.16 0.00 1.01 (0.22)	0.13 0.00 1.02 (0.21)	0.14 0.00 1.02 (0.21)
0.30	0.25 0.00 1.01 (0.27)	0.13 0.00 1.01 (0.21)	0.12 0.00 1.01 (0.18)	0.11 0.00 1.01 (0.17)	0.08 0.00 1.01 (0.17)
0.40	0.21 0.00 1.01 (0.24)	0.12 0.00 1.01 (0.18)	0.09 0.00 1.01 (0.16)	0.06 0.00 1.01 (0.15)	0.05 0.00 1.01 (0.14)
0.50	0.17 0.00 1.00 (0.21)	0.09 0.00 1.01 (0.16)	0.05 0.00 1.01 (0.14)	0.04 0.00 1.01 (0.13)	0.04 0.00 1.01 (0.13)
0.60	0.14 0.00 1.00 (0.19)	0.06 0.00 1.00 (0.15)	0.04 0.00 1.00 (0.13)	0.03 0.00 1.01 (0.12)	0.03 0.00 1.01 (0.12)
0.70	0.12 0.01 1.00 (0.17)	0.06 0.01 1.00 (0.13)	0.03 0.00 1.00 (0.12)	0.02 0.00 1.00 (0.11)	0.02 0.00 1.01 (0.11)
0.80	0.10 0.05 0.94 (0.15)	0.04 0.03 0.99 (0.12)	0.03 0.02 1.00 (0.11)	0.02 0.02 1.00 (0.10)	0.01 0.03 1.01 (0.10)
0.90	0.09 0.24 0.77 (0.19)	0.03 0.14 0.87 (0.14)	0.02 0.12 0.93 (0.10)	0.01 0.00 0.97 (0.09)	0.00 0.00 0.99 (0.09)

Notes. [a] The probability of violating the "four-fifths" rule.
[b] The probability that a "four-fifths" rule violation will be statistically significant (i.e., the ratio is statistically less than .80).
[c] The expected ratio (i.e., mean).
[d] The standard deviation of the expected ratio.

Table 4

Proportion of Time Organization Will Violate Four-Fifths Rule, the Probability of Identifying the Violation as Statistically Significant, and the Mean Minority/Majority Applicant Hiring Ratio and Its Standard Deviation with 5,000 Applicants

Selection Ratio	Minority Group Applicant Base Rate				
	0.10	0.20	0.30	0.40	0.50
0.10	0.08[a] 0.00[b] 1.00[c] (0.15)[d]	0.03 0.00 1.00 (0.11)	0.01 0.00 1.00 (0.10)	0.01 0.00 1.00 (0.09)	0.01 0.00 1.00 (0.09)
0.20	0.02 0.00 1.00 (0.11)	0.00 0.00 1.00 (0.08)	0.00 0.00 1.00 (0.07)	0.00 0.00 1.00 (0.06)	0.00 0.00 1.00 (0.06)
0.30	0.01 0.00 1.00 (0.09)	0.00 0.00 1.00 (0.06)	0.00 0.00 1.00 (0.06)	0.00 0.00 1.00 (0.05)	0.00 0.00 1.00 (0.05)
0.40	0.00 0.00 1.00 (0.07)	0.00 0.00 1.00 (0.06)	0.00 0.00 1.00 (0.05)	0.00 0.00 1.00 (0.05)	0.00 0.00 1.00 (0.04)
0.50	0.00 0.00 1.00 (0.07)	0.00 0.00 1.00 (0.05)	0.00 0.00 1.00 (0.04)	0.00 0.00 1.00 (0.04)	0.00 0.00 1.00 (0.04)
0.60	0.00 0.00 1.00 (0.06)	0.00 0.00 1.00 (0.05)	0.00 0.00 1.00 (0.04)	0.00 0.00 1.00 (0.04)	0.00 0.00 1.00 (0.04)
0.70	0.00 0.00 1.00 (0.06)	0.00 0.00 1.00 (0.04)	0.00 0.00 1.00 (0.04)	0.00 0.00 1.00 (0.03)	0.00 0.00 1.00 (0.03)
0.80	0.00 0.00 1.00 (0.05)	0.00 0.00 1.00 (0.04)	0.00 0.00 1.00 (0.03)	0.00 0.00 1.00 (0.03)	0.00 0.00 1.00 (0.03)
0.90	0.00 0.01 0.99 (0.05)	0.00 0.00 1.00 (0.04)	0.00 0.00 1.00 (0.03)	0.00 0.00 1.00 (0.03)	0.00 0.00 1.00 (0.03)

Notes. [a] The probability of violating the "four-fifths" rule.

[b] The probability that a "four-fifths" rule violation will be statistically significant (i.e., the ratio is statistically less than .80).

[c] The expected ratio (i.e., mean).

[d] The standard deviation of the expected ratio.

variability for most conditions except for those involving high selection ratio with accompanying low minority representation in the applicant pool.

With 5,000 applicants the situation becomes non-volatile. The probability that the "four-fifths" rule would be violated was essentially zero except for selection ratios around 0.10. The probability of finding statistical significance for an individual violation was essentially zero. The expected ratio of the hiring rates for an applicant pool of 5,000 varied little, hovering around 1.00 with low to very low variability.

DISCUSSION

The conversion of inter-group score differences to an indication of whether the "four-fifths" rule is expected to be violated provides human resources decision-makers with the impact information with which they are most concerned. Table 1 enables I/O psychologists to easily convert from the standard deviation inter-group difference metric to an adverse impact metric. Tables 2 through 4 can be used to identify the minimum probability of finding an adverse impact situation for the case where there was no mean group difference on the selection measure.

For typical situations involving a selection ratio of .50 or less, the results clearly showed that small group differences of 0.10 to 0.27 of a standard deviation are all that can be tolerated before one would expect to be faced with an adverse impact problem. When applied to specific applicant samples it is clear that even a mean inter-group difference of 0.10 of a standard deviation would be problematic when the applicant pool is only 100. This can be extrapolated from the fact that adverse impact is problematic with applicant pools of 100, even when there is no inter-group difference on the selection measure. For applicant pools of 500 with no mean inter-group difference, an adverse impact finding is still highly possible, although it is not likely to be substantiated statistically. It is not known what the degree of the problem would be when the mean inter-group difference on the selection measure rose to 0.10 or higher for applicant pools of 500, but the probability of finding adverse impact would be higher as would the probability of statistically substantiating it.

These results illustrate the value of converting mean inter-group differences on a selection measure to the kinds of statistics that directly address the issue of concern, namely, "how likely is it that I will be faced with adverse impact and statistically substantiated adverse impact if I use a selection measure with inter-group differences?" These results indicate (i.e., applicant pool of 100) and suggest (i.e., applicant pool of 500) that it is likely that adverse impact will have to be addressed when even small mean inter-group differences are coupled with small applicant pools.

Assumptions

These estimates were made with various assumptions. Fortunately all but one assumption were reasonable. The one that is not specifies that all selected applicants will accept the employment offer. The extent of the deviation from full offer acceptance, the relationship

between accepting the offer and the selection score, and the relationship between accepting the offer and group membership will all have an impact on the estimates provided herein. This is one area in which more work is needed to identify values for these parameters and to incorporate them into the computations. One reasonable prediction is that higher scorers are more likely to reject the offer, irrespective of group membership. A cursory analysis based on such an assumption indicates that a less than full acceptance rate would increase, not decrease, the expected maximum delta. Under such a scenario the current maximum delta can be viewed as a conservative estimate, and a larger difference would actually have to be observed to yield an expected violation of the "four-fifths" rule. Under this hypothesis the results based on the specific applicant pool sizes would not change since the groups are assumed to have equal means on the selection measure.

Future Research

One issue not yet addressed is that the decision-maker may be concerned with more than two applicant groups. Perhaps the decision-maker is concerned with not violating the "four-fifths" for either Blacks and Hispanics in comparison to Whites (i.e., three groups). In that case, the methodology can easily be modified to incorporate three groups or any other number of groups. The essential elements for the two group case apply to cases involving any number of groups.

Another issue is the sampling fluctuation around the expected maximum delta. The expected maximum delta is always correct only for applicant samples of infinite size. For specific applicant samples of 100, 1,000, or any other size not infinite, the "four-fifths" rule may be violated even though it would not be violated in the population of applicants from which it was drawn. The amount of fluctuation about these expected maximum deltas is not known but it will most likely vary as a function of the number of applicants. It would be useful in future work to also estimate the variability of these expected maximum deltas. This would give the decision-maker the ability to use a conservative estimate of the maximum delta which would apply to a specific size sample of applicants.

A final recommendation is to modify the current methodology to compute the probability of adverse impact at various inter-group differences and applicant pool sizes. This would essentially create a table, much like tables 2, 3, or 4, for each inter-group difference/applicant pool size combination. This would allow for exact estimation of adverse impact for specific inter-group difference/applicant pool size combinations.

REFERENCES

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Wienfeld, F. D., & York, R. L. (1966). Equality of educational opportunity. Washington, DC: US Government Printing Office.

Grant, D. L., & Bray, D. W. (1970). Validation of employment tests for telephone company installation and repair occupations. Journal of Applied Psychology, 54, 7-14.

Hunter, J. E. (1983). Fairness of the General Aptitude Test Battery: Ability differences and their impact on minority hiring rates (U.S. Employment Service Test Research Report No. 46). Washington, DC: Employment and Training Administration. (ERIC Document Reproduction Service No. ED 237 534)

Hunter, J. E., Schmidt, F. L., & Rauschenberger, J. M. (1977). Fairness of psychological tests: Implications of four definitions of selection utility and minority hiring. Journal of Applied Psychology, 62, 245-260.

Hyde, J. S., Fennema, E., & Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. Psychological Bulletin, 107, 139-155.

Hyde, J. S., & Linn, M. C. (1988). Gender differences in verbal ability: A meta-analysis. Psychological Bulletin, 104, 53-69.

Uniform guidelines on employee selection procedures (1978). Federal Register, 43, 38290-38315.